An introduction to network meta-analysis for decision making

Sofia Dias, Nicky Welton, David Phillippo

G-I-N 2018, Manchester

University of BRISTOL

Conflict of Interest and Funding

We have no actual or potential conflict of interest in relation to this worksop.

NICE National Institute for Health and Care Excellence

This work was supported by the Centre for Clinical Practice, NICE, with funding to the Clinical Guidelines Technical Support Unit, University of Bristol.

MRC Medical Research Council

This work was undertaken with the support of the MRC ConDuCT-II Hub (MR/K025643/1), and the MRC grant MR/M005232/1.

Any views expressed are those of the authors and not of the funders.

University of BRISTOL

Network meta-analysis: introduction and assumptions

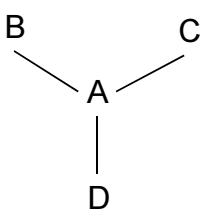
University of BRISTOL

What is a Network Meta-analysis?

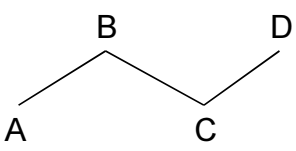
A —— B Pair-wise MA

Indirect Comparisons

(a)

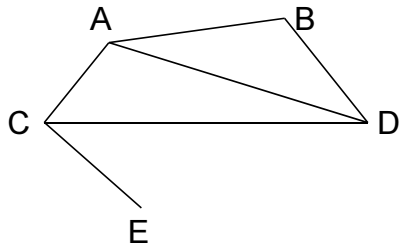


(b)



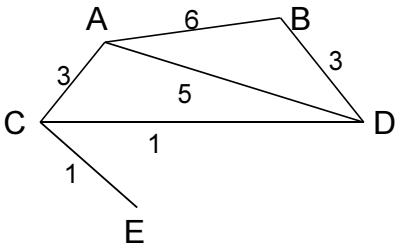
University of BRISTOL

Network Meta-analysis: loops



AKA: Mixed Treatment Comparison (MTC)
Multiple Treatment Comparison
Multiple Treatment Meta-analysis (MTM)

Network Meta-analysis: loops



Number of trials added to edges

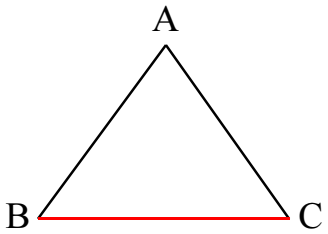
AKA: Mixed Treatment Comparison (MTC)
Multiple Treatment Comparison
Multiple Treatment Meta-analysis (MTM)

Significance of LOOPS in NMA

The existence of "evidence loops" means that there is both DIRECT evidence and INDIRECT evidence on the same contrast

- More data, so estimates are more precise, more robust (less sensitive to any one source of data)
- Possible, and indeed *necessary*, to check "consistency" of the direct and indirect.

Red line represent **DIRECT** evidence on BC effect;
Black lines provide **INDIRECT** evidence on BC effect.



University of BRISTOL

When to do a NMA?

IF

a) There are more than 2 treatments

AND

b) If all the trials had been on just two treatments (**any two**), you would have done a pair-wise MA

University of BRISTOL

Do you HAVE to do a NMA?

If you want to decide which is best out of >2: **YES!**

Suppose 3 treatments: A,B,C

They have been compared in B vs A, C vs A and C vs B trials

Separate meta-analyses have been done and we now have estimates $\hat{d}_{AB}, \hat{d}_{AC}, \hat{d}_{BC}$

Is that enough? **NO!**

We need COHERENT ESTIMATES

$$\hat{d}_{AC}^{COH} = \hat{d}_{AB}^{COH} + \hat{d}_{BC}^{COH}$$

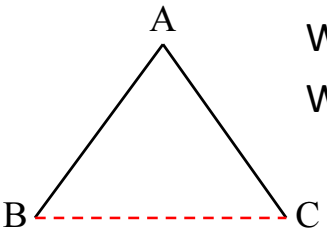
Coherent decisions

- Coherent decisions need coherent evidence
- Otherwise no way to decide which treatment is best
- NMA is just a method to find the coherent estimates

Coherent decisions

- Does it make sense to make decisions based on estimates of relative effect from separate meta-analyses?
  - These estimates will not be consistent, yet we are still using them to make the decision.
- By stating we can use all evidence to make a decision, we are implicitly assuming that the evidence is consistent…
  - And doing informal indirect comparisons without fully accounting for uncertainty

Indirect comparisons: Bucher method



We have "direct" data on AB and AC effects.

What is the BC effect?

$$\hat{d}_{BC}^{indirect} = \hat{d}_{AC}^{direct} - \hat{d}_{AB}^{direct}$$

$$\therefore \text{var}(\hat{d}_{BC}^{indirect}) = \text{var}(\hat{d}_{AC}^{direct}) + \text{var}(\hat{d}_{AB}^{direct})$$

Exercise 1. Indirect Comparisons

Outcome: mean difference (reduction in pain)

Comparison	Mean difference	SE	95%CI
B vs A	-2.3	0.45	(-3.18, -1.42)
C vs A	-4.0	0.50	(-4.98, -3.02)

In the absence of a direct RCT, calculate the relative effect of treatment C vs B

(assuming the patient populations included in the B vs A and the C vs A trials are comparable to each other and relevant to our target population)

Exercise 1. Solution

Comparison	Mean difference	SE	95%CI
B vs A	-2.3	0.45	(-3.18, -1.42)
C vs A	-4.0	0.50	(-4.98, -3.02)

Indirect estimate of C vs B:

$$\hat{d}_{BC}^{Ind} = \hat{d}_{AC}^{Dir} - \hat{d}_{AB}^{Dir} = -4.0 - (-2.3) = -1.7$$

With 95%CI $(-3.01, -0.39)$

Network meta-analysis

Indirect estimate of C vs B:

$$\hat{d}_{BC}^{Ind} = \hat{d}_{AC}^{Dir} - \hat{d}_{AB}^{Dir} = -1.7 \qquad 95\%\text{CI } (-3.01,\ -0.39)$$

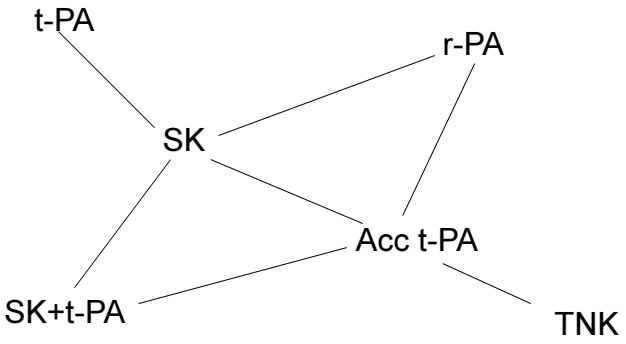
What if we also had a trial of C vs B:
mean difference -1.8 95%CI (-3.66, 0.06)?

- A network meta-analysis will put together direct and indirect evidence to produce a coherent estimate using all the available, relevant, evidence.

Thrombolysis NMA: worked example

Boland et al, *Health Technology Assessment*, 2003

RCTs	SK	t-PA	Acc t-PA	Sk+tPA	r-PA	TNK
8	✓	✓				
1	✓		✓	✓		
1	✓			✓		
1	✓				✓	
2			✓		✓	
1			✓			✓



6 treatments - Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated tissue-plasminogen activator (At-PA), Tenecteplase (TNK), Reteplase (r-PA)
14 trials. **15 possible pairwise comparisons**

Results: Thrombolysis – pairwise meta-analyses (fixed effects) odds ratios (95%CI)

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00** (0.94-1.06)	**0.86** (0.78-0.94)	**0.96** (0.87-1.05)	**0.95** (0.79-1.12)	
t-PA		X				
Acc t-PA			X	**1.12** (1.00-1.25)	**1.02** (0.90-1.16)	**1. 01** (0.88-1.14)
t-PA+SK				X		
r-PA					X	
TNK						X

OR > 1 favours row-defining treatment

Note: all results from fixed effects analysis AND same direction of effects…

University of BRISTOL

Upper right: pair-wise ORs 95%CI
Lower left: NMA ORs 95% CI

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00** (0.94-1.06)	**0.86** (0.78-0.94)	**0.96** (0.87-1.05)	**0.95** (0.79-1.12)	
t-PA	**1.00** (0.94-1.06)	X				
Acc t-PA	**0.87** (0.79-0.94)	**0.87** (0.78-0.97)	X	**1.12** (1.00-1.25)	**1.02** (0.90-1.16)	**1.01** (0.88-1.14)
t-PA+SK	**0.96** (0.88-1.05)	**0.96** (0.87-1.08)	**1.11** (1.00-1.24)	X		
r-PA	**0.90** (0.80-1.01)	**0.90** (0.79-1.03)	**1.04** (0.94-1.16)	**0.94** (0.81-1.08)	X	
TNK	**0.87** (0.75-1.01)	**0.87** (0.74-1.03)	**1.01** (0.89-1.14)	**0.90** (0.77-1.07)	**0.96** (0.82-1.14)	X

Upper diagonal: OR > 1 favours row-defining treatment

Lower diagonal: OR < 1 favours row-defining treatment

University of BRISTOL

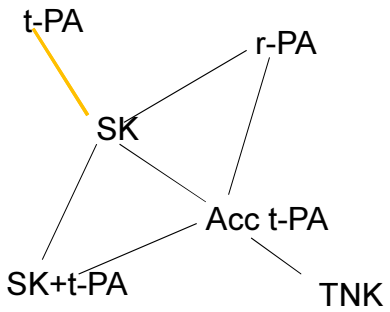
Upper right: pair-wise ORs 95%CI
Lower left: NMA ORs 95% CI

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00** (0.94-1.06)	**0.86** (0.78-0.94)	**0.96** (0.87-1.05)	**0.95** (0.79-1.12)	
t-PA	**1.00** (0.94-1.06)	X				
Acc t-PA	**0.87** (0.79-0.94)	**0.87** (0.78-0.97)	X	**1.12** (1.00-1.25)	**1.02** (0.90-1.16)	**1.01** (0.88-1.14)
t-PA+SK	**0.96** (0.88-1.05)	**0.96** (0.87-1.08)	**1.11** (1.00-1.24)	X		
r-PA	**0.90** (0.80-1.01)	**0.90** (0.79-1.03)	**1.04** (0.94-1.16)	**0.94** (0.81-1.08)	X	
TNK	**0.87** (0.75-1.01)	**0.87** (0.74-1.03)	**1.01** (0.89-1.14)	**0.90** (0.77-1.07)	**0.96** (0.82-1.14)	X

Upper diagonal: OR > 1 favours row-defining treatment
Lower diagonal: OR < 1 favours row-defining treatment

Example: Early thrombolysis for AMI

RCTs	SK	t-PA	Acc t-PA	Sk+tPA	r-PA	TNK
8	✓	✓				
1	✓		✓	✓		
1	✓			✓		
1	✓				✓	
2			✓		✓	
1			✓			✓



6 treatments: Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated t-PA (Acc t-PA), Tenecteplase (TNK), Reteplase (r-PA)
14 trials; **7 comparisons** made; **15 possible** pairwise comparisons

Upper right: pair-wise ORs 95%CI
Lower left: NMA ORs 95% CI

	SK	t-PA	Acc t-PA	t-PA+SK	r-PA	TNK
SK	X	**1.00** (0.94-1.06)	**0.86** (0.78-0.94)	**0.96** (0.87-1.05)	**0.95** (0.79-1.12)	
t-PA	**1.00** (0.94-1.06)	X				
Acc t-PA	**0.87** (0.79-0.94)	**0.87** (0.78-0.97)	X	**1.12** (1.00-1.25)	**1.02** (0.90-1.16)	**1.01** (0.88-1.14)
t-PA+SK	**0.96** (0.88-1.05)	**0.96** (0.87-1.08)	**1.11** (1.00-1.24)	X		
r-PA	**0.90** (0.80-1.01)	**0.90** (0.79-1.03)	**1.04** (0.94-1.16)	**0.94** (0.81-1.08)	X	
TNK	**0.87** (0.75-1.01)	**0.87** (0.74-1.03)	**1.01** (0.89-1.14)	**0.90** (0.77-1.07)	**0.96** (0.82-1.14)	X

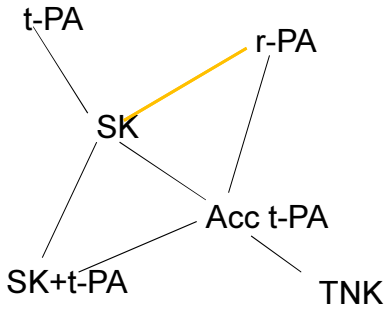
Upper diagonal: OR > 1 favours row-defining treatment
Lower diagonal: OR < 1 favours row-defining treatment

G-I-N 2018: Introduction to NMA for decision making 21

University of BRISTOL

Example: Early thrombolysis for AMI*

RCTs	SK	t-PA	Acc t-PA	Sk+tPA	r-PA	TNK
8	✓	✓				
1	✓		✓	✓		
1	✓			✓		
1	✓				✓	
2			✓		✓	
1			✓			✓



6 treatments: Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated t-PA (Acc t-PA), Tenecteplase (TNK), Reteplase (r-PA)
14 trials; **7 comparisons** made; **15 possible** pairwise comparisons

G-I-N 2018: Introduction to NMA for decision making 22

University of BRISTOL

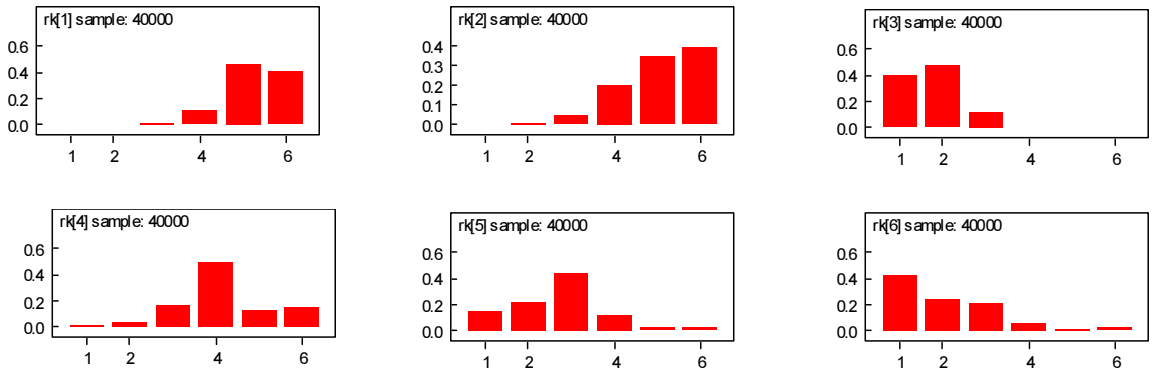
Probability treatment x is 'best': clinical effectiveness

			Ranks		
		Pr(best)	Mean	Median	95% CrI
1	SKA	0	5.27	5	(4, 6)
2	t-PA	0	5.07	5	(3, 6)
3	Acc t-PA	0.40	1.73	2	(1, 3)
4	t-PA + SK	0.01	4.14	4	(2, 6)
5	r-PA	0.16	2.73	3	(1, 5)
6	TNK	0.43	2.06	2	(1, 5)

- **Caution:** Pr(best) very sensitive to different levels of uncertainty.
- It is possible for a treatment to have the **highest probability of being the best** and also the **highest probability of being the worst**!
- Also useful to look at prob of being 2nd best, 3rd best etc

University of BRISTOL

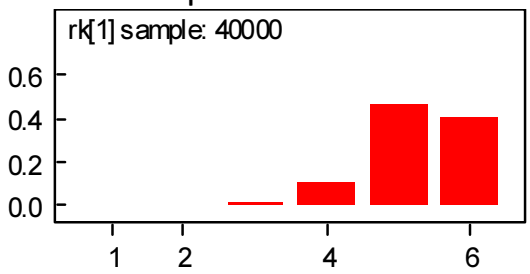
Ranking probability of each treatment



University of BRISTOL

Interpreting ranking plots

- Out of 40,000 iterations treatment 1 was never ranked first, i.e. Pr(treatment 1 is the best)=0
- It was ranked 4th a few times (approx 10%), so Pr(treatment 1 is 4th best) ≈0.10
  - NOTE: this is quite different from: "The prob that treatment 1 is the best 4th line treatment is 10%"



Exercise 2. Interpreting ranking plots

A. Based on the ranking plots in the previous slides, and all other results presented, which thrombolytic treatment (or treatments) would you recommend and why?

(based only on clinical effectiveness)

B. If treatment X is the recommended first line treatment for a particular condition, how would we determine which is the best second line treatment?

Explain your reasons.

Exercise 2. Solution

A. Treatments 3 and 6 have the two highest probabilities of being best and 2nd best, so that overall they both have a high probability of being the top two treatments.

   In addition the odds ratio of treatment 6 compared to 3 is 1.01 95%CrI (0.89-1.14) indicating no difference between them.

   If everything else was equal (costs, side effects etc), we would probably recommend both treatments.

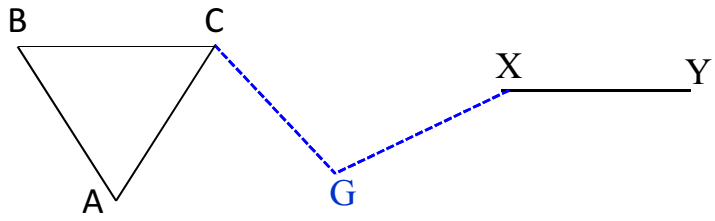
B. To decide which is the best 2nd line therapy, we would need evidence (RCTs) on patients who have failed on the first line treatment, which could be pooled in a NMA to determine the best second line treatment

How wide should a network be?

- Start with the treatments “of interest”
- These should already have been determined in initial discussions / scoping
- Include ALL the trials in the target population between any pair of the treatments of interest.
- Does this form a connected network?
- If so: that is the “base-case” network

Adding a treatment to connect the network

If the network is NOT connected, are there other treatments (G) that can be used to connect the sub-networks ABC and XY?

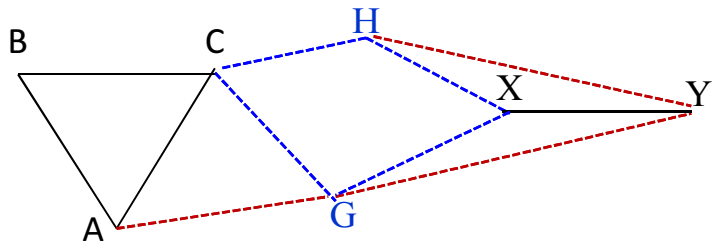


All trials in enlarged network

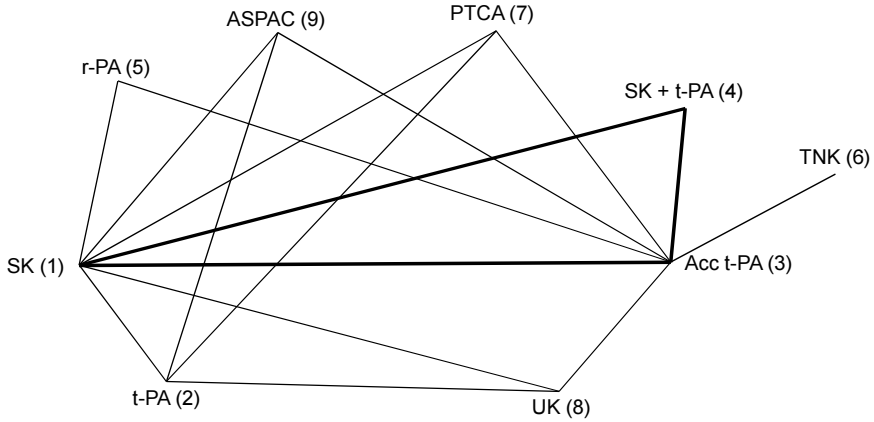
- If an additional treatment is introduced to link networks, check whether there are other treatments that could also do the link, and include.
- Then include ALL the trials that link the enlarged set of treatments

Trial inclusion: multiple connectors

- If an additional treatment is introduced to link networks, check whether there are *other* treatments that could also do the link
- Then include ALL the trials that link the enlarged set of treatments

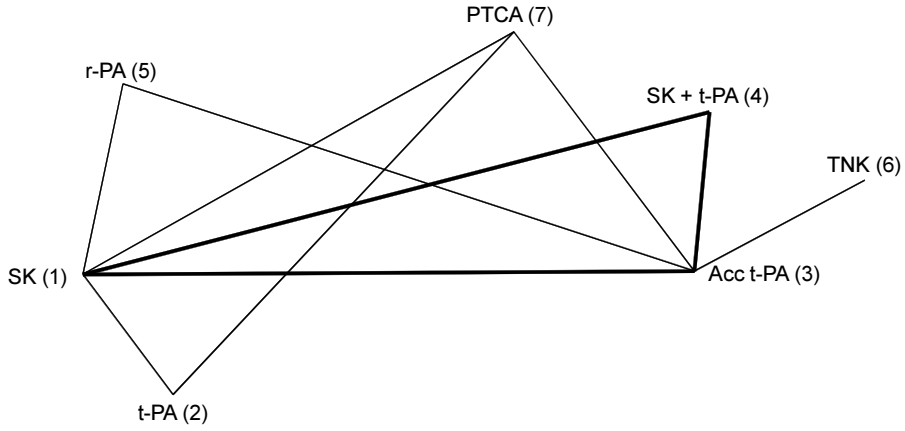


Exercise 3. Extended Thrombolysis Network



- ASPAC, UK, and PTCA were omitted from the Thrombolysis HTA
- ASPAC, UK are old treatments and no longer used. PTCA is used.
- Discuss which evidence you would include in a NMA to inform clinical guidelines

Adding in PTCA



Results: Adding in PTCA (Caldwell, Ades & Higgins; 2005)

	SK	t-PA	Acc t-PA	SK + t-PA	r-PA	TNK	PTCA
SK	**	1.00	0.86	0.96	0.95		0.52
t-PA	1.00	**					0.63
Acc t-PA	0.86	0.86	**	1.12	1.02	1.01	0.81
SK + t-PA	0.96	0.96	1.12	**			
r-PA	0.90	0.90	1.05	0.94	**		
TNK	0.86	0.86	1.01	0.90	0.96	**	
PTCA	0.63	0.64	0.74	0.66	0.71	0.74	**

CI: (0.64 – 1.02)

CI: (0.61 – 0.89)

Probability treatment x is best: clinical effectiveness

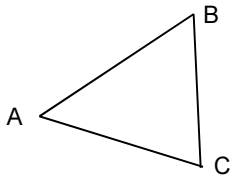
Treatment	% 35 mortality Mean (95%CrI)	Probability Best
SK	6.7 (5.7, 7.7)	0%
t-PA	6.7 (5.7, 7.8)	0%
Acc t-PA	5.8 (4.9, 6.8)	0%
SK + t-PA	6.5 (5.4, 7.6)	0%
R-PA	6.1 (5.0, 7.2)	0%
TNK	5.8 (4.7, 7.1)	0.4%
PTCA	4.4 (3.5, 5.3)	99.6%

(Caldwell, DM. Ades, AE. & Higgins, JPT; *BMJ* 2005)

What have we assumed?

- NMA assumes that the direct and indirect evidence "agree"

$$\theta_{AC} = \theta_{AB} + \theta_{BC}$$

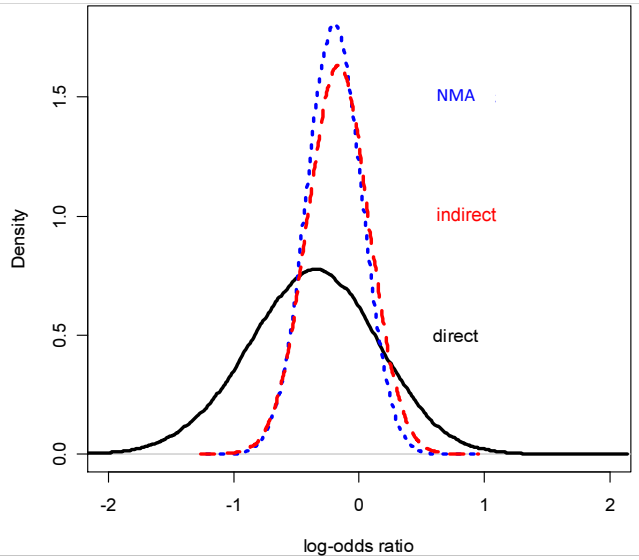


- True effects must be consistent
  - … however data may not be
  - must check for this

$$\theta_{AC} = \theta_{AB} + \theta_{BC}$$

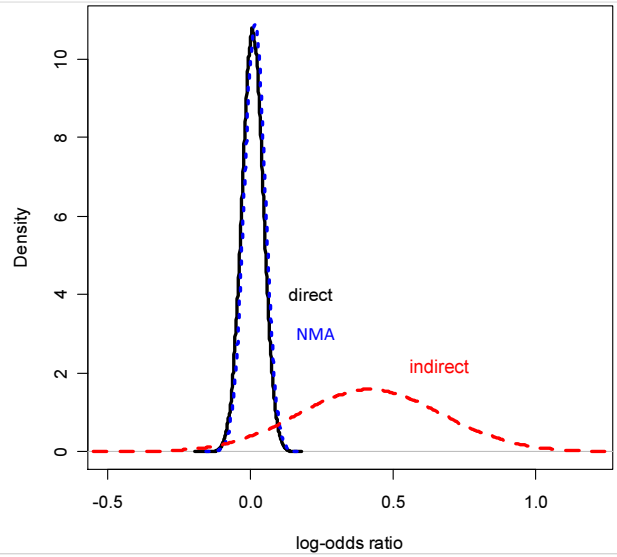
University of BRISTOL

Consistent



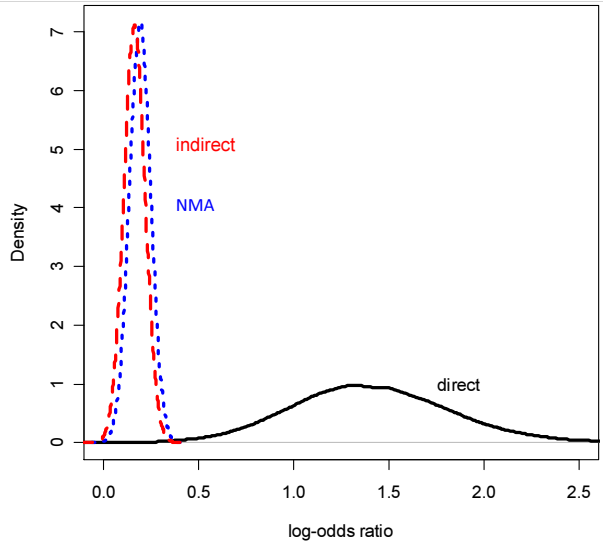
University of BRISTOL

Possibly inconsistent?



University of BRISTOL

Inconsistent



University of BRISTOL

When might consistency assumption fail?

- If trials differ in patients populations & protocols
  - Intervention A differs in AB trials and AC trials e.g. lumping over dose, different controls (waitlist, pill placebo, etc), change in standard care over time …
  - Sicker patients included in AB trials than AC trials
  - Differences in study conduct (outcome measures, risk of bias etc.)
- … ie differences in treatment effect-modifiers
  - Same as heterogeneity in pairwise M-A
- CANNOT check consistency in a simple indirect comparison, but can with NMA if there are loops of evidence

Chou et al, Lancet 2006, 368, 1503-15

Trials of HAART regimes for HIV

A: 2 NRTIs

B: 2 NRTIs + PI

C: 2 NRTIs + NNRTI

"Indirect evidence on d_{BC} inconsistent with direct evidence from BC trials"

"Indirect Comparisons unreliable for complex interventions like HAART"

Dangers of "Lumping"

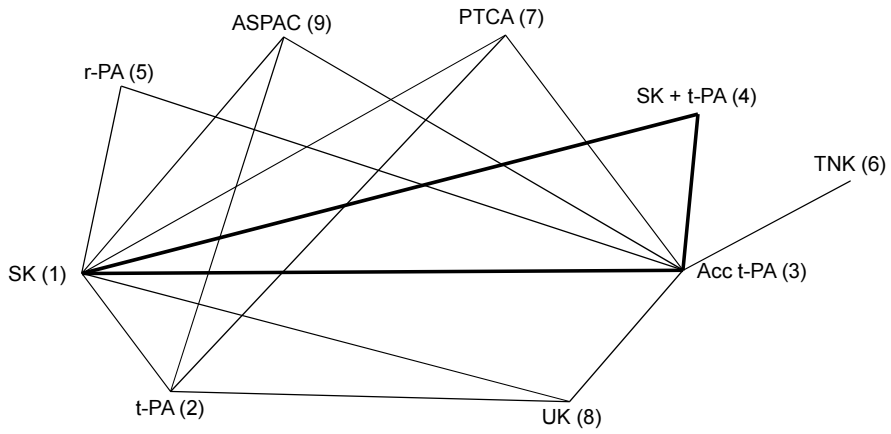
A: 2 NRTIs , B: 2 NRTIs + PI, C: 2 NRTIs + NNRTI

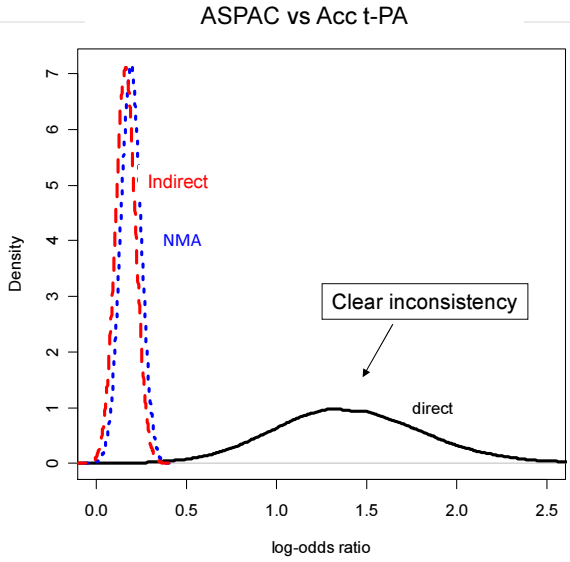
BUT the NRTIs in the AB trials were DIFFERENT from the NRTIs in the BC trials.

When the comparison was restricted to trials with the SAME NRTI regimes, the "inconsistency" no longer statistically significant.

Similar findings reported by Song et al : "lumping" over aspirin doses.

Thrombolysis: Including older treatments ASPAC and UK





- Direct information contradicts indirect
- p-value = 0.001
- Indirect evidence from very large trials dominates direct evidence from smaller trials

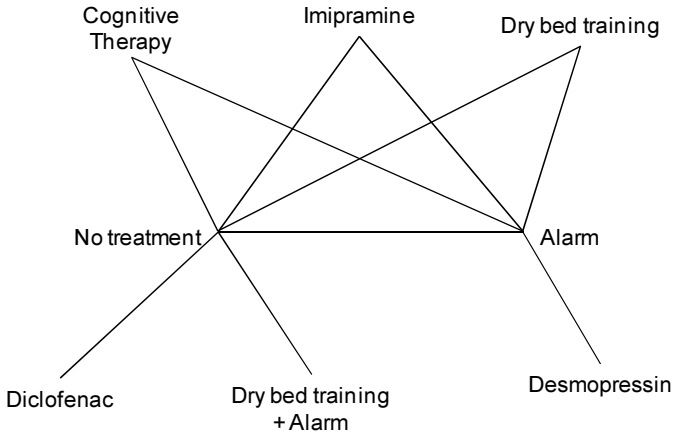
How to avoid inconsistency?

- Careful thought given to PICOS
- Population: well-defined inclusion/exclusion criteria … not too broad
- Interventions … avoid "lumping"
- Comparators … keep different control conditions distinct
- Outcomes … should be similar eg follow-up time, same units
- Study design … exclude studies at high risk of bias

Enuresis Overview of Reviews: Failure to achieve 14 dry nights

X	Y	Relative Risk Y rel to X, θ_{XY}	95% CI
Nothing	Alarm	0.38	0.33 – 0.45
Dry Bed Training	Alarm	1.33	0.79 – 2.24
Desmopressin	Alarm	0.71	0.50 – 0.99
Imipramine	Alarm	0.73	0.61 – 0.88
Cognitive	Alarm	0.68	0.52 – 0.90
Nothing	Cognitive	0.69	0.55 – 0.85
Nothing	Dry Bed Training	0.82	0.66 – 1.02
Nothing	DBT + alarm	0.17	0.11 – 0.28
Nothing	Dicolfenac	0.52	0.38 – 0.70
Nothing	Imipramine	0.77	0.72 – 0.83

Exercise: Enuresis network



- Can the data presented in overview Table lead to a coherent decision, as it stands?

Solution

- No!
  - Alarm vs Nothing RR=0.38
  - Dry Bed Training vs Nothing RR=0.82
    - Suggests Alarm is better than Dry Bed Training
  - Alarm vs Dry Bed Training RR=1.33
    - Suggests Dry Bed Training is better than Alarm
  - Even if direction of effect unchanged, the strength and size of effect can be different

Exercise 4: Enuresis

- Inspect handout with populations and settings for the studies included in the Enuresis example for Alarm vs No Treatment and Imipramine vs No Treatment
- Do you think it is sensible to combine these studies in a NMA?
- What might you do differently?

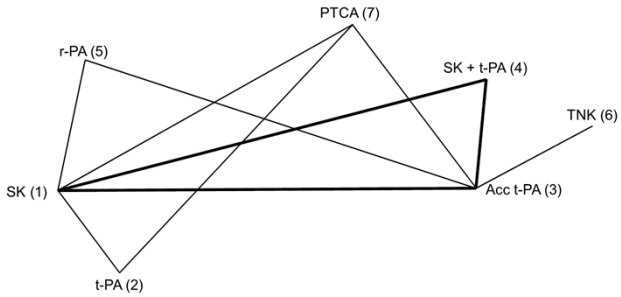
What to do if inconsistency identified?

- Try to explain it …
  - Check the data!
  - Re-visit inclusion/exclusion of studies, intervention definitions and outcomes
  - Consider adjusting for effect modifiers
  - How robust are recommendations to potential bias in the evidence?

Threshold analysis

Assessing the robustness of decisions to changes in the evidence

Motivation – Thrombolytics Example

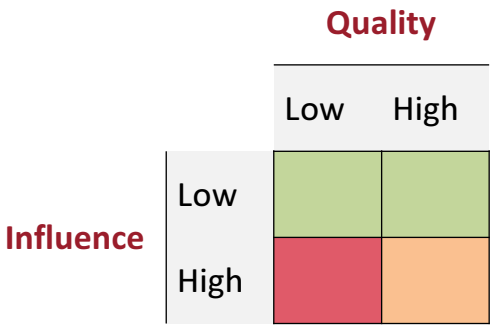


	Treatment	Log odds ratio (95% CrI)
1	SK	0
2	t-PA	0.00 (0.06, 0.06)
3	Accelerated t-PA	-0.16 (-0.24, -0.07)
4	SK plus t-PA	-0.04 (-0.13, 0.05)
5	r-PA	-0.11 (-0.23, 0.01)
6	TNK	-0.15 (-0.30, 0.00)
7	PTCA	-0.46 (-0.66, -0.26)

Motivation

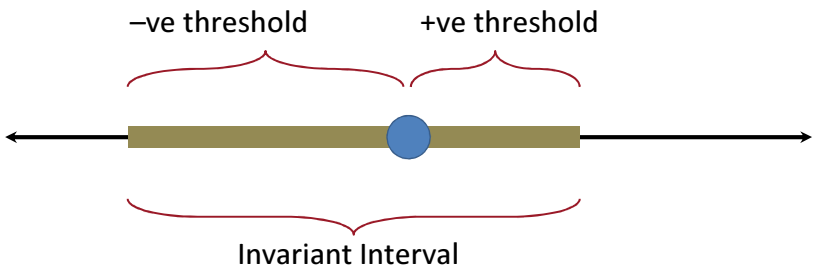
How robust are the results to bias?

- Evidence quality is only half the story



The Threshold Method

We create an *invariant interval* for a data point:

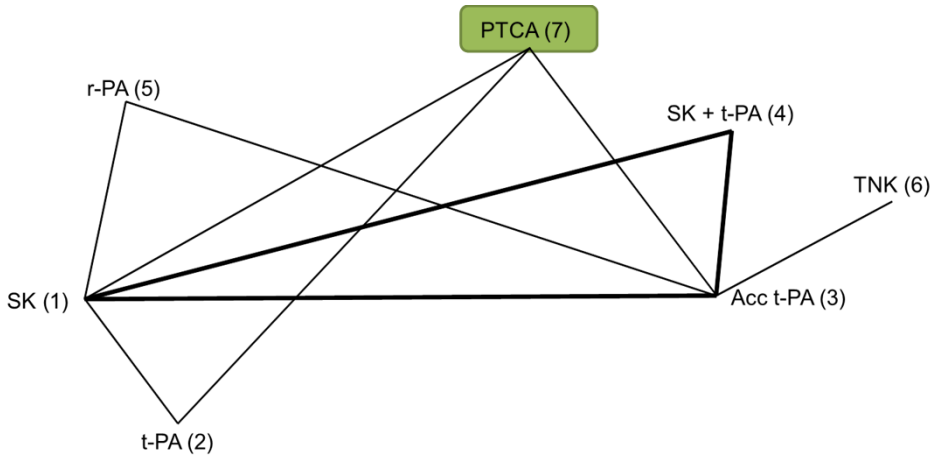


The Threshold Method

We can do this at two "levels":

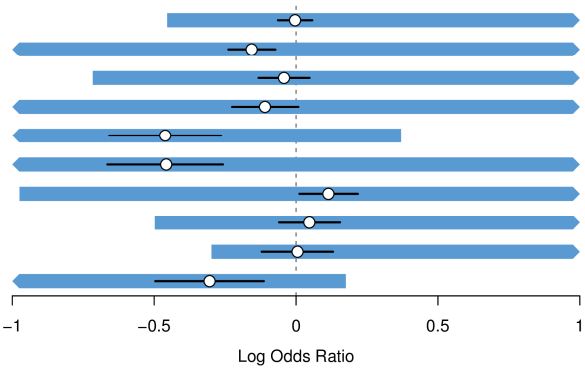
- Study level
  - Thresholds for each individual study estimate
- Contrast level
  - Thresholds for combined body of evidence on a contrast
  - Highly flexible due to an approximation step

Example: Thrombolytics

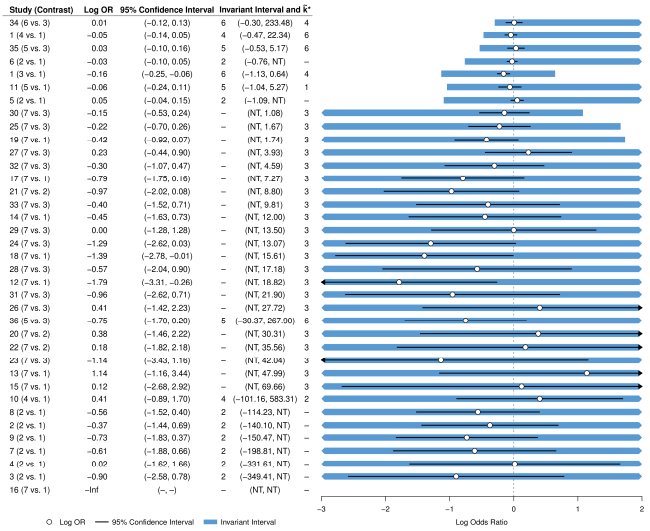


Example: Thrombolytics – contrast level

Contrast	Log OR	95% Credible Interval	Invariant Interval and k̃*		
2 vs. 1	−0.00	(−0.06, 0.06)	2	(−0.45, NT)	–
3 vs. 1	−0.16	(−0.24, −0.07)	3	(−1.43, 1.12)	2
4 vs. 1	−0.04	(−0.13, 0.05)	4	(−0.72, 3.83)	2
5 vs. 1	−0.11	(−0.23, 0.01)	5	(−1.08, 5.17)	1
7 vs. 1	−0.46	(−0.66, −0.26)	–	(NT, 0.37)	3
7 vs. 2	−0.46	(−0.66, −0.26)	–	(NT, NT)	–
4 vs. 3	0.11	(0.01, 0.22)	4	(−0.97, 3.55)	6
5 vs. 3	0.05	(−0.06, 0.15)	5	(−0.50, 4.99)	6
6 vs. 3	0.01	(−0.12, 0.13)	6	(−0.30, 234.30)	4
7 vs. 3	−0.30	(−0.50, −0.11)	–	(NT, 0.18)	3



Example: Thrombolytics – study level

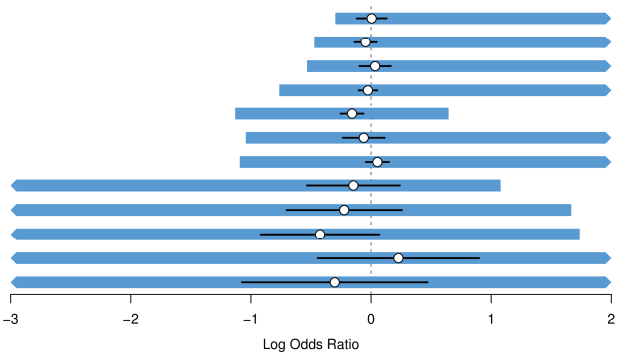


Study (Contrast)	Log OR	95% Confidence Interval		Invariant Interval and k̃*	
34 (6 vs. 3)	0.01	(−0.12, 0.13)	6	(−0.30, 233.48)	4
1 (4 vs. 1)	−0.05	(−0.14, 0.05)	4	(−0.47, 22.34)	6
35 (5 vs. 3)	0.03	(−0.10, 0.16)	5	(−0.53, 5.17)	6
6 (2 vs. 1)	−0.03	(−0.10, 0.05)	2	(−0.76, NT)	–
1 (3 vs. 1)	−0.16	(−0.25, −0.06)	6	(−1.13, 0.64)	4
11 (5 vs. 1)	−0.06	(−0.24, 0.11)	5	(−1.04, 5.27)	1
5 (2 vs. 1)	0.05	(−0.04, 0.15)	2	(−1.09, NT)	–
30 (7 vs. 3)	−0.15	(−0.53, 0.24)	–	(NT, 1.08)	3
25 (7 vs. 3)	−0.22	(−0.70, 0.26)	–	(NT, 1.67)	3
19 (7 vs. 1)	−0.42	(−0.92, 0.07)	–	(NT, 1.74)	3
27 (7 vs. 3)	0.23	(−0.44, 0.90)	–	(NT, 3.93)	3
32 (7 vs. 3)	−0.30	(−1.07, 0.47)	–	(NT, 4.59)	3
17 (7 vs. 1)	−0.79	(−1.75, 0.16)	–	(NT, 7.27)	3
21 (7 vs. 2)	−0.97	(−2.02, 0.08)	–	(NT, 8.80)	3
33 (7 vs. 3)	−0.40	(−1.52, 0.71)	–	(NT, 9.81)	3
14 (7 vs. 1)	−0.45	(−1.63, 0.73)	–	(NT, 12.00)	3
29 (7 vs. 3)	0.00	(−1.28, 1.28)	–	(NT, 13.50)	3
24 (7 vs. 3)	−1.29	(−2.62, 0.03)	–	(NT, 13.07)	3
18 (7 vs. 1)	−1.39	(−2.78, −0.01)	–	(NT, 15.61)	3
28 (7 vs. 3)	−0.57	(−2.04, 0.90)	–	(NT, 17.18)	3
12 (7 vs. 1)	−1.79	(−3.31, −0.26)	–	(NT, 18.82)	3
31 (7 vs. 3)	−0.96	(−2.62, 0.71)	–	(NT, 21.90)	3
26 (7 vs. 3)	0.41	(−1.42, 2.23)	–	(NT, 27.72)	3
36 (5 vs. 3)	−0.75	(−1.70, 0.20)	5	(−30.37, 267.00)	6
20 (7 vs. 2)	0.38	(−1.46, 2.22)	–	(NT, 30.31)	3
22 (7 vs. 2)	0.18	(−1.82, 2.18)	–	(NT, 35.56)	3
23 (7 vs. 3)	−1.14	(−3.43, 1.16)	–	(NT, 42.04)	3
13 (7 vs. 1)	1.14	(−1.16, 3.44)	–	(NT, 47.99)	3
15 (7 vs. 1)	0.12	(−2.68, 2.92)	–	(NT, 69.66)	3
10 (4 vs. 1)	0.41	(−0.89, 1.70)	4	(−101.16, 583.31)	2
8 (2 vs. 1)	−0.56	(−1.52, 0.40)	2	(−114.23, NT)	–
2 (2 vs. 1)	−0.37	(−1.44, 0.69)	2	(−140.10, NT)	–
9 (2 vs. 1)	−0.73	(−1.83, 0.37)	2	(−150.47, NT)	–
7 (2 vs. 1)	−0.61	(−1.88, 0.66)	2	(−198.81, NT)	–
4 (2 vs. 1)	0.02	(−1.62, 1.66)	2	(−331.61, NT)	–
3 (2 vs. 1)	−0.90	(−2.58, 0.78)	2	(−349.41, NT)	–
16 (7 vs. 1)	−Inf	(−, −)	–	(NT, NT)	–

University of BRISTOL

Example: Thrombolytics – study level

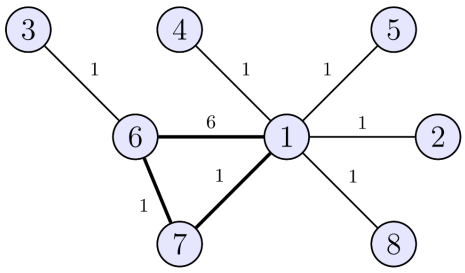
Study (Contrast)	Log OR	95% Confidence Interval		Invariant Interval and k̃*	
34 (6 vs. 3)	0.01	(−0.12, 0.13)	6	(−0.30, 233.48)	4
1 (4 vs. 1)	−0.05	(−0.14, 0.05)	4	(−0.47, 22.34)	6
35 (5 vs. 3)	0.03	(−0.10, 0.16)	5	(−0.53, 5.17)	6
6 (2 vs. 1)	−0.03	(−0.10, 0.05)	2	(−0.76, NT)	–
1 (3 vs. 1)	−0.16	(−0.25, −0.06)	6	(−1.13, 0.64)	4
11 (5 vs. 1)	−0.06	(−0.24, 0.11)	5	(−1.04, 5.27)	1
5 (2 vs. 1)	0.05	(−0.04, 0.15)	2	(−1.09, NT)	–
30 (7 vs. 3)	−0.15	(−0.53, 0.24)	–	(NT, 1.08)	3
25 (7 vs. 3)	−0.22	(−0.70, 0.26)	–	(NT, 1.67)	3
19 (7 vs. 1)	−0.42	(−0.92, 0.07)	–	(NT, 1.74)	3
27 (7 vs. 3)	0.23	(−0.44, 0.90)	–	(NT, 3.93)	3
32 (7 vs. 3)	−0.30	(−1.07, 0.47)	–	(NT, 4.59)	3



Log OR — 95% Confidence Interval — Invariant Interval

University of BRISTOL

Exercise: Headaches clinical guideline



	Treatment	Mean change in headache days per month (95% CrI)
1	Placebo	0
2	Telmisartan	-0.51 (-2.32, 1.27)
3	Amitriptyline	-1.14 (-2.45, 0.16)
4	Divalproex Sodium	0.13 (-0.99, 1.23)
5	Gabapentin	0.00 (-1.60, 1.58)
6	Topiramate	-1.04 (-1.52, -0.58)
7	Propranolol	-1.19 (-2.20, -0.20)
8	Propranolol/Nadolol	-0.60 (-1.65, 0.45)

Minimal Clinically Important Difference (MCID) = 0.5 days per month

(NICE CG151.1, 2015)

Exercise 5. Headaches clinical guideline

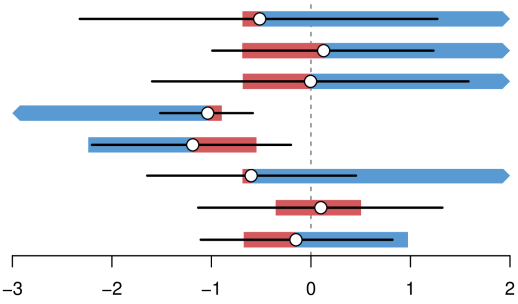
For the contrast and study level plots in turn:

- Are there any contrasts/studies that the decision is not sensitive to?
- Are there any that the decision is sensitive to?

What do you conclude about the robustness of the treatment decision?

Exercise 5. Headaches clinical guideline – contrast level

Contrast	Mean	95% CrI	Invariant Interval and k̃*		
2 vs. 1	−0.51	(−2.32, 1.27)	2, 3, 6, 7	(−0.69, NT)	–
4 vs. 1	0.13	(−0.99, 1.23)	3, 4, 6, 7	(−0.69, 12.02)	3, 6, 7, 8
5 vs. 1	−0.00	(−1.60, 1.58)	3, 5, 6, 7	(−0.69, 24.04)	2, 3, 6, 7
6 vs. 1	−1.04	(−1.51, −0.58)	–	(NT, −0.90)	3, 6, 7, 8
7 vs. 1	−1.19	(−2.20, −0.20)	3, 7	(−2.24, −0.55)	3, 6, 7, 8
8 vs. 1	−0.60	(−1.64, 0.45)	3, 6, 7, 8	(−0.69, NT)	–
6 vs. 3	0.10	(−1.13, 1.32)	6, 7	(−0.35, 0.50)	3, 7
7 vs. 6	−0.15	(−1.11, 0.82)	3, 7	(−0.67, 0.97)	3, 6



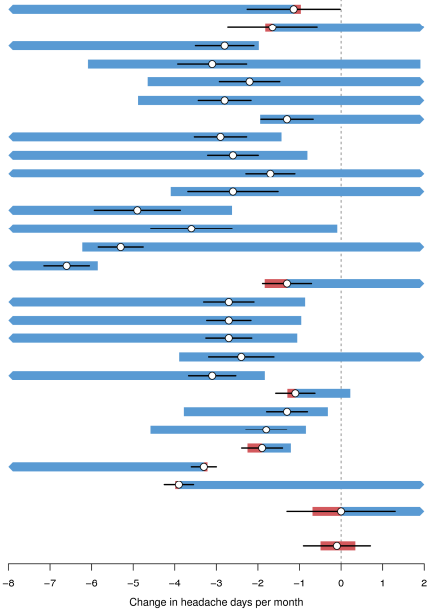
○ Mean — 95% CrI Invariant Interval

Base−case optimal treatment set is 3, 6, 7.

University of BRISTOL

Exercise 5. Headaches clinical guideline – study level

Study (Trt.)	Mean	95% CI	Invariant Interval and k̃*		
Diener 2009 (1)	−1.14	(−2.26, −0.02)	–	(NT, −0.97)	2, 3, 6, 7
Diener 2009 (2)	−1.65	(−2.72, −0.58)	2, 3, 6, 7	(−1.82, NT)	–
Apostol 2008 (1)	−2.80	(−3.50, −2.10)	3, 6, 7, 8	(−9.95, −1.98)	3, 4, 6, 7
Apostol 2008 (4)	−3.10	(−3.93, −2.27)	3, 4, 6, 7	(−6.09, 1.91)	3, 6, 7, 8
Apostol 2008 (4)	−2.20	(−2.93, −1.47)	3, 4, 6, 7	(−4.65, NT)	–
Apostol 2008 (4)	−2.80	(−3.43, −2.17)	3, 4, 6, 7	(−4.88, 5.19)	3, 6, 7, 8
Brandes 2004 (1)	−1.30	(−1.93, −0.67)	3, 6, 7, 8	(−1.95, NT)	–
Brandes 2004 (6)	−2.90	(−3.53, −2.27)	–	(NT, −1.43)	3, 6, 7, 8
Brandes 2004 (6)	−2.60	(−3.21, −1.99)	–	(NT, −0.81)	3, 6, 7, 8
Brandes 2004 (6)	−1.70	(−2.29, −1.11)	–	(NT, 15.76)	2, 3, 6, 7
Lewis 2009 (1)	−2.60	(−3.68, −1.52)	3, 6, 7, 8	(−4.10, NT)	–
Lewis 2009 (6)	−4.90	(−5.93, −3.87)	–	(NT, −2.62)	3, 6, 7, 8
Lewis 2009 (6)	−3.60	(−4.57, −2.63)	–	(NT, −0.09)	3, 6, 7, 8
Lipton 2011 (1)	−5.30	(−5.84, −4.76)	3, 6, 7, 8	(−6.22, NT)	–
Lipton 2011 (6)	−6.60	(−7.14, −6.06)	–	(NT, −5.85)	3, 6, 7, 8
Silberstein 2004 (1)	−1.30	(−1.89, −0.71)	3, 6, 7, 8	(−1.84, NT)	–
Silberstein 2004 (6)	−2.70	(−3.30, −2.10)	–	(NT, −0.87)	3, 6, 7, 8
Silberstein 2004 (6)	−2.70	(−3.23, −2.17)	–	(NT, −0.96)	3, 6, 7, 8
Silberstein 2004 (6)	−2.70	(−3.25, −2.15)	–	(NT, −1.05)	3, 6, 7, 8
Winner 2005 (1)	−2.40	(−3.18, −1.62)	3, 6, 7, 8	(−3.89, NT)	–
Winner 2005 (6)	−3.10	(−3.67, −2.53)	–	(NT, −1.83)	3, 6, 7, 8
Diener 2004 (1)	−1.10	(−1.57, −0.63)	3, 6, 7, 8	(−1.29, 0.22)	3, 7
Diener 2004 (6)	−1.30	(−1.79, −0.81)	3, 6	(−3.78, −0.32)	3, 7
Diener 2004 (6)	−1.60	(−2.29, −1.31)	3, 6	(−4.56, −0.65)	3, 7
Diener 2004 (7)	−1.90	(−2.39, −1.41)	3, 7	(−2.25, −1.21)	3, 6
Holroyd 2010 (1)	−3.30	(−3.60, −3.00)	6, 7	(−17.43, −3.21)	3, 6, 7, 8
Holroyd 2010 (8)	−3.90	(−4.25, −3.55)	3, 6, 7, 8	(−3.99, 10.77)	6, 7
Silberstein 2013 (5 vs. 1)	0.00	(−1.30, 1.30)	3, 5, 6, 7	(−0.69, 11.76)	2, 3, 6, 7
Dodick 2009 (3 vs. 6)	−0.10	(−0.90, 0.70)	3, 7	(−0.49, 0.34)	6, 7



○ Mean — 95% CI Invariant Interval

Base−case optimal treatment set is 3, 6, 7.

Conclusions

- Threshold analysis provides insight into the effects of bias adjustment on treatment decisions
- We can have more confidence in recommendations where thresholds are large
- We can focus attention on the quality of decision-sensitive trials and contrasts
- More complex analyses can investigate specific concerns in the evidence, e.g. in groups of studies or treatments
- Can be used with a range of decision rules or for decisions based on cost-effectiveness

References

- Boland A, Dundar Y, Bagust A, Haycox A, Hill R, Mujica Mota R, Walley T, Dickson R. Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. HTA 2003 7:15
- Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005, 331; 897-900.
- Chou et al, Lancet 2006, 368, 1503-15
- Dias, S., A. E. Ades, N. J. Welton, J. P. Jansen and A. J. Sutton (2018). Network meta-analysis for decision making, Wiley.
- Phillippo DM, Dias S, Ades AE, Didelez V and Welton NJ (2018). *Sensitivity of treatment recommendations to bias in network meta-analysis*. J. R. Stat. Soc. A, 181: 843-867. doi:10.1111/rssa.12341
  - R package *nmathresh* package available on CRAN